1. Prove that $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$:

$$\begin{aligned}
\boldsymbol{a} \in \mathcal{C}(\boldsymbol{X}) &\iff \underset{n\times 1}{\boldsymbol{a}} = \boldsymbol{X}\boldsymbol{b} && \text{for some } \underset{p\times 1}{\boldsymbol{b}} \\
&\iff \boldsymbol{a} = \boldsymbol{X} \underset{p\times p}{\boldsymbol{I}} \boldsymbol{b} && \text{for some } \boldsymbol{b}\text{x} \\
&\iff \boldsymbol{a} = \boldsymbol{X}\boldsymbol{B}^{-1} \underbrace{\boldsymbol{B}\boldsymbol{b}}_{p\times 1} && \text{treat as } \boldsymbol{X}\boldsymbol{B}^{-1} \text{ product a } p\times 1 \text{ vector} \\
&\implies \boldsymbol{a} \in \mathcal{C}(\boldsymbol{P}_{\boldsymbol{X}})
\end{aligned}$$

So $\mathcal{C}(\boldsymbol{X}) \subseteq \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$.
Then similarly,

$$\begin{aligned}
\boldsymbol{g} \in \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1}) &\iff \boldsymbol{g} = \boldsymbol{X}\boldsymbol{B}^{-1}\boldsymbol{h} && \text{for some } p\times 1 \text{ vector } \boldsymbol{h} \\
&\iff \boldsymbol{g} = \boldsymbol{X} \underbrace{\boldsymbol{B}^{-1}\boldsymbol{h}}_{p\times 1} && \text{treat as } \boldsymbol{X} \text{ product a } p\times 1 \text{ vector} \\
&\implies \boldsymbol{g} \in \mathcal{C}(\boldsymbol{X})
\end{aligned}$$

So $\mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1}) \subseteq \mathcal{C}(\boldsymbol{X})$.
According to the results above, $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$.

2. (a) Rewrite these hypotheses in terms of the $\boldsymbol{\alpha}$ parameter vector.

$$H_0 : \boldsymbol{\alpha}_2 = \mathbf{0} \text{ vs. } H_A : \boldsymbol{\alpha}_2 \neq \mathbf{0}$$

(b) If we want to fit a reduced model corresponding to the null hypothesis, we could use $\boldsymbol{W}_1$ because $\boldsymbol{W}\boldsymbol{\alpha} = \boldsymbol{W}_1\boldsymbol{\alpha}_1 + \boldsymbol{W}_2\boldsymbol{\alpha}_2$ reduces down to $\boldsymbol{W}_1\boldsymbol{\alpha}_1$ under the null hypothesis $H_0 : \boldsymbol{\alpha}_2 = \mathbf{0}$.

(c) Provide a matrix $\boldsymbol{C}$ for testing the main effect of temperature.
The main effect of temperature is

$$\frac{1}{2}(\mu_{11} + \mu_{21}) - \frac{1}{2}(\mu_{12} + \mu_{22}) = \left(\frac{1}{2}, \quad -\frac{1}{2}, \quad \frac{1}{2}, \quad -\frac{1}{2}\right) \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{pmatrix}$$

So one possible matrix $\boldsymbol{C}$ is $\left(\frac{1}{2}, \quad -\frac{1}{2}, \quad \frac{1}{2}, \quad -\frac{1}{2}\right)$.

(d) Provide a matrix $\boldsymbol{A}$ so that the rank of $\boldsymbol{B}$ is 4. Many answers are possible. One choice that works is

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

(e) Provide a model matrix for a reduced model that corresponds to the null hypothesis of no temperature main effect.

$$\boldsymbol{W} = \boldsymbol{X}\boldsymbol{B}^{-1} = \begin{bmatrix} \underset{2\times 1}{\mathbf{1}} & & & \\ & \underset{3\times 1}{\mathbf{1}} & & \\ & & \underset{4\times 1}{\mathbf{1}} & \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \underset{2\times 1}{\mathbf{1}} & & & \\ & \underset{3\times 1}{\mathbf{1}} & & \\ & & \underset{4\times 1}{\mathbf{1}} & \\ 1 & -1 & 1 & -2 \end{bmatrix}$$

By part (b), the model matrix of the reduced model is the first $p - q = 3$ columns of $\boldsymbol{W}$, i.e.,

$$\boldsymbol{W}_1 = \begin{bmatrix} \underset{2\times 1}{\mathbf{1}} & & \\ & \underset{3\times 1}{\mathbf{1}} & \\ & & \underset{4\times 1}{\mathbf{1}} \\ 1 & -1 & 1 \end{bmatrix}.$$

(f) Find the error sum of squares for the reduced and full models.
$SS_{Reduced} = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}_1})\boldsymbol{y} = 185.28$
$SS_{Full} = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}})\boldsymbol{y} = 12$

(g) Find the degrees of freedom associated with the sums of squares in part (f).
$df_{Reduced} = rank(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}_1}) = 10 - 3 = 7$
$df_{Full} = rank(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}}) = 10 - 4 = 6$

(h) Compute the F -statistic for testing the null hypothesis of no time main effect using the sums of squares and degrees of freedom computed in parts (f) and (g).
From the code and output below,

$$F = \frac{(SSE_{Reduced} - SSE_{Full})/(df_{Reduced} - df_{Full})}{SSE_{Full}/df_{Full}} = 86.64$$

```
> X=model.matrix(~ 0 + factor(c(rep(1, 2), rep(2, 3), rep(3, 4), 4)))
> B=rbind(cbind(diag(3), 0), .5 * c(1, -1, 1, -1))
> W=X %*% solve(B)
> W1=W[, -4]
> proj=function(x) {
+    x %*% MASS::ginv(t(x) %*% x) %*% t(x)
+ }
> Pw=proj(W)
> Pw1=proj(W1)
> y=c(3, 5, 11, 13, 15, 5, 6, 6, 7, 16)
> I10=diag(rep(1, length(y)))
> SS.red=t(y) %*% (I10 - Pw1) %*% y
> SS.full=t(y) %*%(I10 - Pw) %*% y
> (SS.red - SS.full) / SS.full * 6
```

3. We can conduct a lack of fit F-test for this problem. The full model is a cell mean model with 4 parameters, and the reduced model has 2 parameters with one for the combination of diet 1 with drug 1 and one for all the other three combinations. Since we have 40 observations, the degree of freedom for the full model is $df_{full} = 40 - 4 = 36$, the degree of freedom for the reduced model is $df_{reduced} = 40 - 2 = 38$, and the difference $df_{reduced} - df_{full} = 38 - 26 = 2$. So the F-distribution we would use is $F_{2,36}$.

   We have the noncentrality parameter as the following form:

$$\frac{(\boldsymbol{C\beta} - \boldsymbol{d})'[\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}']^{-1}(\boldsymbol{C\beta} - \boldsymbol{d})}{2\sigma^2}$$

   where we define:

$$\boldsymbol{\beta} = \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{pmatrix} \quad \boldsymbol{C} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \quad \boldsymbol{d} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{10\times1} & 0 & 0 & 0 \\ 0 & \mathbf{1}_{10\times1} & 0 & 0 \\ 0 & 0 & \mathbf{1}_{10\times1} & 0 \\ 0 & 0 & 0 & \mathbf{1}_{10\times1} \end{bmatrix}$$

   So we can get:

$$\boldsymbol{C} - \boldsymbol{d} = \begin{pmatrix} \mu_{12} - \mu_{21} \\ \mu_{12} - \mu_{22} \end{pmatrix}$$

   and:

$$[\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}']^{-1} = \frac{10}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

   Then the noncentrality parameter is:

$$\begin{aligned} &\frac{\begin{pmatrix} \mu_{12} - \mu_{21} \\ \mu_{12} - \mu_{22} \end{pmatrix}' \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{pmatrix} \mu_{12} - \mu_{21} \\ \mu_{12} - \mu_{22} \end{pmatrix}}{2\sigma^2} \times \frac{10}{3} \\ =&\frac{10}{3\sigma^2} \times \left[(\mu_{12} - \mu_{21})^2 + (\mu_{12} - \mu_{22})^2 - (\mu_{12} - \mu_{21})(\mu_{12} - \mu_{22})\right] \end{aligned}$$

   An alternative method could also be used here as following and we get an equivalent expression. The null hypothesis can be tested with a F-test comparing a reduced model to a full model. Let

$$\boldsymbol{X}_0 = \begin{bmatrix} \mathbf{1}_{10\times1} & 0 \\ 0 & \mathbf{1}_{30\times1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{10\times1} & 0 & 0 & 0 \\ 0 & \mathbf{1}_{10\times1} & 0 & 0 \\ 0 & 0 & \mathbf{1}_{10\times1} & 0 \\ 0 & 0 & 0 & \mathbf{1}_{10\times1} \end{bmatrix}$$

Also $r_0 = 2, r = 4$.

From slide set 3, we know the F statistic has a noncentral F distribution with degree of freedom $r - r_0$ and $n - r$. In this problem, this's $4 - 2 = 2$ and $40 - 4 = 36$, then

$$\begin{aligned}\text{NCP} &= \frac{\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{P}_X - \boldsymbol{P}_{X_0})\boldsymbol{X}\boldsymbol{\beta}}{2\sigma^2} \\ &= ||(\boldsymbol{P}_X - \boldsymbol{P}_{X_0})\boldsymbol{X}\boldsymbol{\beta}||^2/(2\sigma^2) \\ &= ||\boldsymbol{P}_X\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{P}_{X_0}\boldsymbol{X}\boldsymbol{\beta}||^2/(2\sigma^2) \\ &= ||\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{P}_{X_0}\boldsymbol{X}\boldsymbol{\beta}||^2/(2\sigma^2)\end{aligned}$$

Now note that:

$$\boldsymbol{X}\boldsymbol{\beta} = \begin{bmatrix} \mu_{11}\mathbf{1}_{10\times 1} \\ \mu_{12}\mathbf{1}_{10\times 1} \\ \mu_{21}\mathbf{1}_{10\times 1} \\ \mu_{22}\mathbf{1}_{10\times 1} \end{bmatrix}$$

$$\begin{aligned}\boldsymbol{P}_{X_0}\boldsymbol{X}\boldsymbol{\beta} &= \boldsymbol{X}_0(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{X}_0'\boldsymbol{X}\boldsymbol{\beta} \\ &= \boldsymbol{X}_0(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{X}_0' \begin{bmatrix} \mu_{11}\mathbf{1}_{10\times 1} \\ \mu_{12}\mathbf{1}_{10\times 1} \\ \mu_{21}\mathbf{1}_{10\times 1} \\ \mu_{22}\mathbf{1}_{10\times 1} \end{bmatrix} \\ &= \boldsymbol{X}_0 \begin{bmatrix} 1/10 & 0 \\ 0 & 1/30 \end{bmatrix} \begin{bmatrix} 10\mu_{11} \\ 10(\mu_{12} + \mu_{21} + \mu_{22}) \end{bmatrix} \\ &= \boldsymbol{X}_0 \begin{bmatrix} \mu_{11} \\ \bar{\mu} \end{bmatrix}, \quad \text{where } \ \bar{\mu} \equiv \frac{\mu_{12} + \mu_{21} + \mu_{22}}{3} \\ &= \begin{bmatrix} \mu_{11}\mathbf{1}_{10\times 1} \\ \bar{\mu}\mathbf{1}_{30\times 1} \end{bmatrix}\end{aligned}$$

Thus,

$$\begin{aligned}&||\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{P}_{X_0}\boldsymbol{X}\boldsymbol{\beta}||^2 \\ =&\left|\left| \begin{bmatrix} \mu_{11}\mathbf{1}_{10\times 1} \\ \mu_{12}\mathbf{1}_{10\times 1} \\ \mu_{21}\mathbf{1}_{10\times 1} \\ \mu_{22}\mathbf{1}_{10\times 1} \end{bmatrix} - \begin{bmatrix} \mu_{11}\mathbf{1}_{10\times 1} \\ \bar{\mu}\mathbf{1}_{30\times 1} \end{bmatrix} \right|\right|^2 \\ =&10[(\mu_{12} - \bar{\mu})^2 + (\mu_{21} - \bar{\mu})^2 + (\mu_{22} - \bar{\mu})^2]\end{aligned}$$

So, the noncentrality parameter is:

$$\begin{aligned}&\frac{10[(\mu_{12} - \bar{\mu})^2 + (\mu_{21} - \bar{\mu})^2 + (\mu_{22} - \bar{\mu})^2]}{2\sigma^2} \\ =&\frac{5[(\mu_{12} - \bar{\mu})^2 + (\mu_{21} - \bar{\mu})^2 + (\mu_{22} - \bar{\mu})^2]}{\sigma^2}\end{aligned}$$